

## SEQUENCE LISTING

## (1) GENERAL INFORMATION:

(i) APPLICANT: MELKI, JUDITH  
MUNNICH, ARNOLD

(ii) TITLE OF INVENTION: SURVIVAL MOTOR NEURON (SMN) GENE: A GENE  
FOR SPINAL MUSCULAR ATROPHY

(iii) NUMBER OF SEQUENCES: 57

(iv) CORRESPONDENCE ADDRESS:

- (A) ADDRESSEE: BIRCH, STEWART, KOLASCH AND BIRCH, LLP  
(B) STREET: PO BOX 747  
(C) CITY: FALLS CHURCH  
(D) STATE: VA  
(E) COUNTRY: USA  
(F) ZIP: 22040-0747

(v) COMPUTER READABLE FORM:

- (A) MEDIUM TYPE: Floppy disk  
(B) COMPUTER: IBM PC compatible  
(C) OPERATING SYSTEM: PC-DOS/MS-DOS  
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:

- (A) APPLICATION NUMBER:  
(B) FILING DATE:  
(C) CLASSIFICATION:

(viii) ATTORNEY/AGENT INFORMATION:

- (A) NAME: FARACI, C. J.  
(B) REGISTRATION NUMBER: 32,350  
(C) REFERENCE/DOCKET NUMBER: 2121-110P

(ix) TELECOMMUNICATION INFORMATION:

- (A) TELEPHONE: (703) 205-8000  
(B) TELEFAX: (703) 205-8050

## (2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

AATTTTAA	TTTTTGTAG	AGACAGGGTC	TCATTATGTT	CCCCAGGGTG	GTGTCAAGCT	60
CCAGGTCTCA	AGTGATCCCC	CTACCTCCGC	CTCCCAAAGT	TGTGGGATTG	TAGGCATGAG	120
CCACTGCAAG	AAAACCTTAA	CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAA	180
GTACATTAAA	AGACTATCAA	CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	240
ATGTCTTGTG	AACAAAATGC	TTTTTAACAT	CCATATAAAG	CTATCTATAT	ATAGCTATCT	300
ATGTCTATAT	AGCTATTTTT	TTTAACTTCC	TTTTATTTTC	CTTACAG		347

## (2) INFORMATION FOR SEQ ID NO:2:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 444 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

GTAGTCTGC	CAGCATTATG	AAAGTGAATC	TTACTTTTGT	AAAACTTTAT	GGTTTGTGGA	60
AAACAAATGT	TTTTGAACAG	TTAAAAAGTT	CAGATGTTAA	AAAGTTGAAA	GGTTAATGTA	120
AAACAATCAA	TATTAAAGAA	TTTTGATGCC	AAAAC TATTA	GATAAAAGGT	TAATCTACAT	180
CCCTACTAGA	ATTCTCATAC	TTAACTGGTT	GGTTATGTGG	AAGAAACATA	CTTTCACAAT	240
AAAGAGCTTT	AGGATATGAT	GCCATTTTAT	ATCACTAGTA	GGCAGACCAG	CAGACTTTTT	300
TTTATTGTGA	TATGGGATAA	CCTAGGCATA	CTGCACTGTA	CACTCTGACA	TATGAAGTGC	360
TCTAGTCAAG	TTTAACTGGT	GTCCACAGAG	GACATGGTTT	AACTGGAATT	CGTCAAGCCT	420
CTGGTTCTAA	TTTCTCATTT	GCAG				444

## (2) INFORMATION FOR SEQ ID NO:3:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 347 base pairs
- (B) TYPE: nucleic acid

(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

AATTTTAAA	TTTTTGTAG	AGACAGGGTC	TCATTATGTT	GCCCAGGGTG	GTGTCAAGCT	60
CCAGGTCTCA	AGTGATCCCC	CTACCTCCGC	CTCCCAAAGT	TGTGGGATTG	TAGGCATGAG	120
CCACTGCAAG	AAAACCTTAA	CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAAA	180
GTACATTAAA	AGACTATCAA	CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	240
ATGCTTGTG	AACAAAATGC	TTTTTAACAT	CCATATAAAG	CTATCTATAT	ATAGCTATCT	300
ATATCTATAT	AGCTATTTTT	TTTAACTTCC	TTTTATTTTC	CTTACAG		347

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 444 base pairs  
(B) TYPE: nucleic acid  
(C) STRANDEDNESS: double  
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTAAGTCTGC	CAGCATTATG	AAAGTGAATC	TTACTTTTGT	AAAACCTTAT	GGTTTGTGGA	60
AAACAAATGT	TTTTGAACAG	TTAAAAAGTT	CAGATGTTAG	AAAGTTGAAA	GGTTAATGTA	120
AAACAATCAA	TATTAAAGAA	TTTTGATGCC	AAAACCTATTA	GATAAAAGGT	TAATCTACAT	180
CCCTACTAGA	ATTCTCATAC	TTAACTGGTT	GGTGTGTGG	AAGAAACATA	CTTTCACAAT	240
AAAGAGCTTT	AGGATATGAT	GCCATTTTAT	ATCACTAGTA	GGCAGACCAG	CAGACTTTTT	300
TTTATTGTGA	TATGGGATAA	CCTAGGCATA	CTGCACTGTA	CACTCTGACA	TATGAAGTGC	360
TCTAGTCAAG	TTTAACTGGT	GTCCACAGAG	GACATGGTTT	AACTGGAATT	CGTCAAGCCT	420
CTGGTTCTAA	TTTCTCATTT	GCAG				444

## (2) INFORMATION FOR SEQ ID NO:5:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

AGACTATCAA CTTAATTCT GATCA

25

## (2) INFORMATION FOR SEQ ID NO:6:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

TAAGGAATGT GAGCACCTTC CTTC

24

## (2) INFORMATION FOR SEQ ID NO:7:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GTAATAACCA AATGCAATGT GAA

23

## (2) INFORMATION FOR SEQ ID NO:8:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CTAGAACACC CTTCTCACAG

20

## (2) INFORMATION FOR SEQ ID NO:9:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 294 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Met	Ala	Met	Ser	Ser	Gly	Gly	Ser	Gly	Gly	Gly	Val	Pro	Glu	Gln	Glu	1	5	10	15
Asp	Ser	Val	Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	Ser	Asp	20	25	30	
Ile	Trp	Asp	Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	Val	Ala	35	40	45	
Ser	Phe	Lys	His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	Ser	Gly	50	55	60	
Lys	Pro	Lys	Thr	Thr	Pro	Lys	Arg	Lys	Pro	Ala	Lys	Lys	Asn	Lys	Ser	65	70	75	80
Gln	Lys	Lys	Asn	Thr	Ala	Ala	Ser	Leu	Gln	Gln	Trp	Lys	Val	Gly	Asp	85	90	95	

Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala Thr  
 100 105 110  
 Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr Thr  
 115 120 125  
 Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser Pro  
 130 135 140  
 Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu Asn Glu  
 145 150 155 160  
 Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg Ser Pro  
 165 170 175  
 Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp Asn Ser  
 180 185 190  
 Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly Pro Gly  
 195 200 205  
 Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro  
 210 215 220  
 Pro Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro  
 225 230 235 240  
 Pro Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp  
 245 250 255  
 Ala Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr  
 260 265 270  
 His Thr Gly Tyr Tyr Met Gly Phe Arg Gln Asn Gln Lys Glu Gly Arg  
 275 280 285  
 Cys Ser His Ser Leu Asn  
 290

## (2) INFORMATION FOR SEQ ID NO:10:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGGGGCCCCA	CGCTGCGCAC	CCGCGGGTTT	GCTATGGCGA	TGAGCAGCGG	CGGCAGTGGT	60
GGCGGCGTCC	CGGAGCAGGA	GGATTCCGTG	CTGTTCCGGC	GCGGCACAGG	CCAGAGCGAT	120
GATTCTGACA	TTTGGGATGA	TACAGCACTG	ATAAAAGCAT	ATGATAAAGC	TGTGGCTTCA	180
TTTAAGCATG	CTCTAAAGAA	TGGTGACATT	TGTGAAACTT	CGGGTAAACC	AAAAACCACA	240
CCTAAAAGAA	AACCTGCTAA	GAAGAATAAA	AGCCAAAAGA	AGAATACTGC	AGCTTCCTTA	300
CAACAGTGGA	AAGTTGGGGA	CAAATGTTCT	GCCATTTGGT	CAGAAGACGG	TTGCATTTAC	360
CCAGCTACCA	TTGCTTCAAT	TGATTTTAAG	AGAGAAACCT	GTGTTGTGGT	TTACACTGGA	420
TATGGAAATA	GAGAGGAGCA	AAATCTGTCC	GATCTACTTT	CCCCAATCTG	TGAAGTAGCT	480
AATAATATAG	AACAGAATGC	TCAAGAGAAT	GAAAATGAAA	GCCAAGTTTC	AACAGATGAA	540
AGTGAGAACT	CCAGGTCTCC	TGGAAATAAA	TCAGATAACA	TCAAGCCCAA	ATCTGCTCCA	600
TGGAAACCCT	TTCTCCCTCC	ACCACCCCCC	ATGCCAGGGC	CAAGACTGGG	ACCAGGAAAG	660
CCAGGTCTAA	AATTCAATGG	CCCACCACCG	CCACCGCCAC	CACCACCACC	CCACTTACTA	720
TCATGCTGGC	TGCCTCCATT	TCCTTCTGGA	CCACCAATAA	TTCCCCCACC	ACCTCCCATA	780
TGTCCAGATT	CTCTTGATGA	TGCTGATGCT	TTGGGAAGTA	TGTTAATTTT	ATGGTACATG	840
AGTGGCTATC	ATACTGGCTA	TTATATGGGT	TTTAGACAAA	ATCAAAAAGA	AGGAAGGTGC	900
TCACATTCCT	TAAATTAAGG	AGAAATGCTG	GCATAGAGCA	GCACTAAATG	ACACCACTAA	960
AGAAACGATC	AGACAGATCT	GGAATGTGAA	GCGTTATAGA	AGATAACTGG	CCTCATTTCT	1020
TCAAAATATC	AAGTGTTGGG	AAAGAAAAAA	GGAAGTGGAA	TGGGTAACTC	TTCTTGATTA	1080
AAAGTTATGT	AATAACCAAA	TGCAATGTGA	AATATTTTAC	TGGACTCTTT	TGAAAAACCA	1140
TCTGTAAAAG	ACTGAGGTGG	GGGTGGGAGG	CCAGCACGGT	GGTGAGGCAG	TTGAGAAAAT	1200
TTGAATGTGG	ATTAGATTTT	GAATGATATT	GGATAATTAT	TGGTAATTTT	ATGGCCTGTG	1260
AGAAGGGTGT	TGTAGTTTAT	AAAAGACTGT	CTTAATTTGC	ATACTTAAGC	ATTTAGGAAT	1320
GAAGTGTTAG	AGTGTCTTAA	AATGTTTCAA	ATGGTTTAAC	AAAATGTATG	TGAGGCGTAT	1380
GTGGCAAAAT	GTTACAGAAT	CTAACTGGTG	GACATGGCTG	TTCATTGTAC	TGTTTTTTTC	1440
TATCTTCTAT	ATGTTTAAAA	GTATATAATA	AAAATATTTA	ATTTTTTTTT	AAAAAAAAAA	1500

AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA 1560

AAAAAAAAAA AAAAAAAAAA AA 1582

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

AATTTTAAA	TTTTTGTAG	AGACAGGGTC	TCATTATGTT	GCCCAGGGTG	GTGTCAAGCT	60
CCAGGTCTCA	AGTGATCCCC	CTACCTCCGC	CTCCCAAAGT	TGTGGGATTG	TAGGCATGAG	120
CCACTGCAAG	AAAACCTTAA	CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAAA	180
GTACATTAAA	AGACTATCAA	CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	240
ATGTCTTGTG	AACAAAATGC	TTTTTAACAT	CCATATAAAG	CTATCTATAT	ATAGCTATCT	300
ATATCTATAT	AGCTATTTTT	TTTAACTTCC	TTTTATTTTC	CTTACAGGGT	TTTAGACAAA	360
ATCAAAAAGA	AGGAAGGTGC	TCACATTCCT	TAAATTAAGG	AGTAAGTCTG	CCAGCATTAT	420
GAAAGTGAAT	CTTACTTTTG	TAAAACTTTA	TGGTTTGTGG	AAAACAAATG	TTTTTGAACA	480
GTTAAAAAGT	TCAGATGTTA	GAAAGTTGAA	AGGTTAATGT	AAAACAATCA	ATATTAAAGA	540
ATTTTGATGC	CAAAACTATT	AGATAAAAGG	TTAATCTACA	TCCCTACTAG	AATTCTCATA	600
CTTAACTGGT	TGGTTGTGTG	GAAGAAACAT	ACTTTCACAA	TAAAGAGCTT	TAGGATATGA	660
TGCCATTTTA	TATCACTAGT	AGGCAGACCA	GCAGACTTTT	TTTTATTGTG	ATATGGGATA	720
ACCTAGGCAT	ACTGCACTGT	ACACTCTGAC	ATATGAAGTG	CTCTAGTCAA	GTTTAACTGG	780
TGTCCACAGA	GGACATGGTT	TAACTGGAAT	TCGTCAAGCC	TCTGGTTCTA	ATTTCTCATT	840
TGCAGGAAAT	GCTGGCATAG	AGCAGCACTA	AATGACACCA	CTAAAGAAAC	GATCAGACAG	900
ATCTGGAATG	TGAAGCGTTA	TAGAAGATAA	CTGGCCTCAT	TTCTTCAAAA	TATCAAGTGT	960
TGGGAAAGAA	AAAAGGAAGT	GGAATGGGTA	ACTCTTCTTG	ATTAAAAGTT	ATGTAATAAC	1020



CAAATGCAAT GTGAAATATT TTACTGGACT CTTTTGAAAA ACCATCTGTA AAAGACTGAG	1080
GTGGGGGTGG GAGGCCAGCA CGGTGGTGAG GCAGTTGAGA AAATTTGAAT GTGGATTAGA	1140
TTTTGAATGA TATTGGATAA TTATTGGTAA TTTTATGGCC TGTGAGAAGG GTGTTGTAGT	1200
TTATAAAAGA CTGTCTTAAT TTGCATACTT AAGCATTTAG GAATGAAGTG TTAGAGTGTC	1260
TTAAAATGTT TCAAATGGTT TAACAAAATG TATGTGAGGC GTATGTGGCA AAATGTTACA	1320
GAATCTAACT GGTGGACATG GCTGTTCACT GTACTGTTTT TTTCTATCTT CTATATGTTT	1380
AAAAGTATAT AATAAAAATA TTTAATTT	1408

## (2) INFORMATION FOR SEQ ID NO:12:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1582 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: double  
 (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: cDNA

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGGGGCCCCA CGCTGCGCAT CCGCGGGTTT GCTATGGCGA TGAGCAGCGG CGGCAGTGGT	60
GGCGGCGTCC CGGAGCAGGA GGATTCCGTG CTGTTCCGGC GCGGCACAGG CCAGAGCGAT	120
GATTCTGACA TTTGGGATGA TACAGCACTG ATAAAAGCAT ATGATAAAGC TGTGGCTTCA	180
TTTAAGCATG CTCTAAAGAA TGGTGACATT TGTGAAACTT CGGGTAAACC AAAAACCACA	240
CCTAAAAGAA AACCTGCTAA GAAGAATAAA AGCCAAAAGA AGAATACTGC AGCTTCCTTA	300
CAACAGTGGA AAGTTGGGGA CAAATGTTCT GCCATTTGGT CAGAAGACGG TTGCATTTAC	360
CCAGCTACCA TTGCTTCAAT TGATTTTAAG AGAGAAACCT GTGTTGTGGT TTACACTGGA	420
TATGGAAATA GAGAGGAGCA AAATCTGTCC GATCTACTTT CCCCAATCTG TGAAGTAGCT	480
AATAATATAG AACAGAATGC TCAAGAGAAT GAAAATGAAA GCCAAGTTTC AACAGATGAA	540
AGTGAGAACT CCAGGTCTCC TGGAAATAAA TCAGATAACA TCAAGCCCAA ATCTGCTCCA	600
TGGAACCTCTT TTCTCCCTCC ACCACCCCCC ATGCCAGGGC CAAGACTGGG ACCAGGAAAG	660
CCAGGTCTAA AATTCAATGG CCCACCACCG CCACCGCCAC CACCACCACC CCACTTACTA	720

TCATGCTGGC TGCCTCCATT TCCTTCTGGA CCACCAATAA TTCCCCCACC ACCTCCCATA	780
TGTCCAGATT CTCTTGATGA TGCTGATGCT TTGGGAAGTA TGTTAATTTT ATGGTACATG	840
AGTGGCTATC ATACTGGCTA TTATATGGGT TTCAGACAAA ATCAAAAAGA AGGAAGGTGC	900
TCACATTCCT TAAATTAAGG AGAAATGCTG GCATAGAGCA GCACTAAATG ACACCACTAA	960
AGAAACGATC AGACAGATCT GGAATGTGAA GCGTTATAGA AGATAACTGG CCTCATTTCT	1020
TCAAAATATC AAGTGTGGG AAAGAAAAA GGAAGTGGAA TGGGTAATC TTCTTGATTA	1080
AAAGTTATGT AATAACCAA TGCAATGTGA AATATTTTAC TGGACTCTTT TGAAAAACCA	1140
TCTGTAAAAG ACTGGGGTGG GGGTGGGAGG CCAGCACGGT GGTGAGGCAG TTGAGAAAAT	1200
TTGAATGTGG ATTAGATTTT GAATGATATT GGATAATTAT TGGTAATTTT ATGGCCTGTG	1260
AGAAAGGGTGT TGTAAGTTTAT AAAAGACTGT CTTAATTTGC ATACTTAAGC ATTTAGGAAT	1320
GAAGTGTTAG AGTGTCTTAA AATGTTTCAA ATGGTTTAAAC AAAATGTATG TGAGGCGTAT	1380
GTGGCAAAAT GTTACAGAAT CTAAGTGGTG GACATGGCTG TTCATTGTAC TGTTTTTTTTC	1440
TATCTTCTAT ATGTTTAAAA GTATATAATA AAAATATTTA ATTTTTTTTTT AAAAAAAAAA	1500
AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA AAAAAAAAAA	1560
AAAAAAAAAA AAAAAAAAAA AA	1582

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1408 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

AATTTTAAAA TTTTGTAG AGACAGGGTC TCATTATGTT GCCCAGGGTG GTGTCAAGCT	60
CCAGGTCTCA AGTGATCCCC CTACCTCCGC CTCCCAAAGT TGTGGGATTG TAGGCATGAG	120
CCACTGCAAG AAAACCTTAA CTGCAGCCTA ATAATTGTTT TCTTTGGGAT AACTTTTAAA	180
GTACATTAAA AGACTATCAA CTTAATTTCT GATCATATTT TGTTGAATAA AATAAGTAAA	240

ATGTCTTGTG AACAAAATGC TTTTAAACAT CCATATAAAG CTATCTATAT ATAGCTATCT	300
ATGTCTATAT AGCTATTTTT TTTAACTTCC TTTTATTTTC CTTACAGGGT TTCAGACAAA	360
ATCAAAAAGA AGGAAGGTGC TCACATTCCT TAAATTAAGG AGTAAGTCTG CCAGCATTAT	420
GAAAGTGAAT CTTACTTTTG TAAAACTTTA TGGTTTGTGG AAAACAAATG TTTTGAACA	480
GTTAAAAAGT TCAGATGTTA AAAAGTTGAA AGGTAAATGT AAAACAATCA ATATTAAAGA	540
ATTTTGATGC CAAAACATTT AGATAAAAGG TTAATCTACA TCCCTACTAG AATTCTCATA	600
CTTAACTGGT TGGTTATGTG GAAGAAACAT ACTTTCACAA TAAAGAGCTT TAGGATATGA	660
TGCCATTTTA TATCACTAGT AGGCAGACCA GCAGACTTTT TTTTATTGTG ATATGGGATA	720
ACCTAGGCAT ACTGCACTGT ACACTCTGAC ATATGAAGTG CTCTAGTCAA GTTTAACTGG	780
TGTCACAGA GGACATGGTT TAACTGGAAT TCGTCAAGCC TCTGGTTCTA ATTTCTCATT	840
TGCAGGAAAT GCTGGCATAG AGCAGCACTA AATGACACCA CTAAAGAAAC GATCAGACAG	900
ATCTGGAATG TGAAGCGTTA TAGAAGATAA CTGGCCTCAT TTCTTCAAAA TATCAAGTGT	960
TGGGAAAGAA AAAAGGAAGT GGAATGGGTA ACTCTTCTTG ATTAAAAGTT ATGTAATAAC	1020
CAAATGCAAT GTGAAATATT TTAAGTGGT CTTTGAATA ACCATCTGTA AAAGACTGGG	1080
GTGGGGGTGG GAGGCCAGCA CGGTGGTGAG GCAGTTGAGA AAATTTGAAT GTGGATTAGA	1140
TTTGAATGA TATTGGATAA TTATTGGTAA TTTTATGGCC TGTGAGAAGG GTGTTGTAGT	1200
TTATAAAGA CTGTCTTAAT TTGCATACTT AAGCATTTAG GAATGAAGTG TTAGAGTGTC	1260
TTAAATGTT TCAAATGGTT TAACAAAATG TATGTGAGGC GTATGTGGCA AAATGTTACA	1320
GAATCTAACT GGTGGACATG GCTGTTTATT GTACTGTTTT TTTCTATCTT CTATATGTTT	1380
AAAAGTATAT AATAAAAATA TTTAATTT	1408

## (2) INFORMATION FOR SEQ ID NO:14:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 360 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

ACCTGACCCA GAGGTCAAGG CTGCAGTGAG ACGAGATTGC CCACTGCCCT CCACCCTGGG	60
TGATAAGAGT GGGACCCTGT TCAAAACATA CACACACACA CACACACACA CACACACACA	120
CACACACACA CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCTCTCT CTCTCAAAAA	180
CACTTGGTCT GTTATTTTTC GAAATTGTCA GTCATAGTTA TCTGTTAGAC CAAAGCTGGT	240
AAGACATTTA TTACATTGCC TCCTACAAC TCAATCAGCTA ATGTATTTGC TATATAGCAA	300
TTACATATGG ATATATTATC TTAGGGGATG GCCAGTATAA AACTGTCACT GAGGAAAGGA	360

## (2) INFORMATION FOR SEQ ID NO:15:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CTTCCACCT AGCCTCCCCA GTAGCTAGGA CTATAGGCGT GCCCACCAAG CTCAGCTATT	60
TTTTATTTAG TAGAGACGGG GTTTCGGCAG CTTAGGCCTC GTTCGAACTC CAGTGTGTGT	120
GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTGTGTGT GTGTAGATAT TTATTCCCCC	180
TCCCCCTTGG AAAAGTAAGT AAGCTCCTAC TAGGAATTTA AAACCTGCTT GATCTATATA	240
AAGACAAACA AGGAAAGACA AACATGGGGG CAGGAAGGAA GGCAGATC	288

## (2) INFORMATION FOR SEQ ID NO:16:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 141 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TCGAGGTAGA TTTGTATTAT ATCCCATGTA CACACACACA CACACACACA CACACACACA	60
CACACACAGA CTTAATCTGT TTACAGAAAT AAAAGGAATA AAATACCGTT TCTACTATAC	120
ACCAAAACTA GCCATCTTGA C	141

## (2) INFORMATION FOR SEQ ID NO:17:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 305 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCCTGAGAAG GCTTCCTCCT GAGTATGCAT AAACATTAC AGCTTGCATG CGTGTGTGTG	60
TGTGTGTGTG TGTGTATGTT TGCTTGCACT GTAAAAACAA TTGCAACATC AACAGAAATA	120
AAATATTAAAG GAATAATTCT CCTCCGACTC TGCCGTTCCA TCCAGTGAAA CTCTTCATTC	180
TGGGGTAAAG TTCCTTCAGT TCTTTCATAG ATAGGTATAT ACTTCATAAG TCAAACAATC	240
AGGCTGGGTG CAGTAGCTCA TGCCTGTAAT CCCAGCCCTT TGGGAGGCCG AGCTGGGCAG	300
ATCGA	305

## (2) INFORMATION FOR SEQ ID NO:18:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 341 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

- (A) DESCRIPTION: /desc = "SYNTHETIC DNA"

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

```

TCCACCCGCC TTGGCCTCCC AAAGCCTGGG ATTACAGGCG TGA CTGCCGC ACCCAGCTGT      60
AAACTGGTTT AATGGTAGAT TTTAGGTATT AACAATAGAT AAAAAGATAC TTTTGGCATA      120
CTGTGTATTG GGATGGGGTT AGAACAGGTG TCTACCCAAG ACATT TACTT AAAATCGCCC      180
TCGAAATGCT ATGTGAGCTG TGTGTGTGTG TGTGTGTGTG TGTGTATTAA GGAAAAGCAT      240
GAAAGTATTT ATGCTTGATT TTTTTTTTTC CTCATAGCTT CATAGTGGAC AGATACATAG      300
TCTAAATCAA AATGTTTAAA CTTTTTATGT CACTTGCTGT C                          341

```

## (2) INFORMATION FOR SEQ ID NO:19:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 278 amino acids
- (B) TYPE: amino acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

```

Met Ala Met Ser Ser Gly Gly Ser Gly Gly Gly Val Pro Glu Gln Glu
1          5          10          15
Asp Ser Val Leu Phe Arg Arg Gly Thr Gly Gln Ser Asp Asp Ser Asp
20          25          30
Ile Trp Asp Asp Thr Ala Leu Ile Lys Ala Tyr Asp Lys Ala Val Ala
35          40          45
Ser Phe Lys His Ala Leu Lys Asn Gly Asp Ile Cys Glu Thr Ser Gly
50          55          60
Lys Pro Lys Thr Thr Pro Lys Arg Lys Pro Ala Lys Lys Asn Lys Ser
65          70          75          80
Gln Lys Lys Asn Thr Ala Ala Ser Leu Gln Gln Trp Lys Val Gly Asp
85          90          95
Lys Cys Ser Ala Ile Trp Ser Glu Asp Gly Cys Ile Tyr Pro Ala Thr
100          105          110
Ile Ala Ser Ile Asp Phe Lys Arg Glu Thr Cys Val Val Val Tyr Thr
115          120          125

```

Gly Tyr Gly Asn Arg Glu Glu Gln Asn Leu Ser Asp Leu Leu Ser Pro  
 130 135 140  
 Ile Cys Glu Val Ala Asn Asn Ile Glu Gln Asn Ala Gln Glu Asn Glu  
 145 150 155 160  
 Asn Glu Ser Gln Val Ser Thr Asp Glu Ser Glu Asn Ser Arg Ser Pro  
 165 170 175  
 Gly Asn Lys Ser Asp Asn Ile Lys Pro Lys Ser Ala Pro Trp Asn Ser  
 180 185 190  
 Phe Leu Pro Pro Pro Pro Pro Met Pro Gly Pro Arg Leu Gly Pro Gly  
 195 200 205  
 Lys Pro Gly Leu Lys Phe Asn Gly Pro Pro Pro Pro Pro Pro Pro Pro  
 210 215 220  
 Pro Pro His Leu Leu Ser Cys Trp Leu Pro Pro Phe Pro Ser Gly Pro  
 225 230 235 240  
 Pro Ile Ile Pro Pro Pro Pro Pro Ile Cys Pro Asp Ser Leu Asp Asp  
 245 250 255  
 Ala Asp Ala Leu Gly Ser Met Leu Ile Ser Trp Tyr Met Ser Gly Tyr  
 260 265 270  
 His Thr Gly Tyr Tyr Met  
 275

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 885 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:

- (A) NAME/KEY: CDS
- (B) LOCATION: 18..881

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

CGGCGTGGTA GCAGGCC ATG GCG ATG GGC AGT GGC GGA GCG GGC TCC GAG  
 Met Ala Met Gly Ser Gly Gly Ala Gly Ser Glu  
 1 5 10

CAG	GAA	GAT	ACG	GTG	CTG	TTC	CGG	CGT	GGC	ACC	GGC	CAG	AGT	GAT	GAT	98
Gln	Glu	Asp	Thr	Val	Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	
			15					20					25			
TCT	GAC	ATT	TGG	GAT	GAT	ACA	GCA	TTG	ATA	AAA	GCT	TAT	GAT	AAA	GCT	146
Ser	Asp	Ile	Trp	Asp	Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	
		30					35					40				
GTG	GCT	TCC	TTT	AAG	CAT	GCT	CTA	AAG	AAC	GGT	GAC	ATT	TGT	GAA	ACT	194
Val	Ala	Ser	Phe	Lys	His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	
	45					50					55					
CCA	GAT	AAG	CCA	AAA	GGC	ACA	GCC	AGA	AGA	AAA	CCT	GCC	AAG	AAG	AAT	242
Pro	Asp	Lys	Pro	Lys	Gly	Thr	Ala	Arg	Arg	Lys	Pro	Ala	Lys	Lys	Asn	
	60				65					70					75	
AAA	AGC	CAA	AAG	AAG	AAT	GCC	ACA	ACT	CCC	TTG	AAA	CAG	TGG	AAA	GTT	290
Lys	Ser	Gln	Lys	Lys	Asn	Ala	Thr	Thr	Pro	Leu	Lys	Gln	Trp	Lys	Val	
				80					85					90		
GGT	GAC	AAG	TGT	TCT	GCT	GTT	TGG	TCA	GAA	GAC	GGC	TGC	ATT	TAC	CCA	338
Gly	Asp	Lys	Cys	Ser	Ala	Val	Trp	Ser	Glu	Asp	Gly	Cys	Ile	Tyr	Pro	
			95				100						105			
GCT	ACT	ATT	ACG	TCC	ATT	GAC	TTT	AAG	AGA	GAA	ACC	TGT	GTC	GTG	GTT	386
Ala	Thr	Ile	Thr	Ser	Ile	Asp	Phe	Lys	Arg	Glu	Thr	Cys	Val	Val	Val	
		110					115					120				
TAT	ACT	GGA	TAT	GGA	AAC	AGA	GAG	GAG	CAA	AAC	TTA	TCT	GAC	CTA	CTT	434
Tyr	Thr	Gly	Tyr	Gly	Asn	Arg	Glu	Glu	Gln	Asn	Leu	Ser	Asp	Leu	Leu	
	125					130					135					
TCC	CCG	ACC	TGT	GAA	GTA	GCT	AAT	AGT	ACA	GAA	CAG	AAC	ACT	CAG	GAG	482
Ser	Pro	Thr	Cys	Glu	Val	Ala	Asn	Ser	Thr	Glu	Gln	Asn	Thr	Gln	Glu	
	140				145					150					155	
AAT	GAA	AGT	CAA	GTT	TCC	ACA	GAC	GAC	AGT	GAA	CAC	TCC	TCC	AGA	TCG	530
Asn	Glu	Ser	Gln	Val	Ser	Thr	Asp	Asp	Ser	Glu	His	Ser	Ser	Arg	Ser	
				160				165						170		
CTC	AGA	AGT	AAA	GCA	CAC	AGC	AAG	TCC	AAA	GCT	GCT	CCG	TGG	ACC	TCA	578
Leu	Arg	Ser	Lys	Ala	His	Ser	Lys	Ser	Lys	Ala	Ala	Pro	Trp	Thr	Ser	
			175					180					185			
TTT	CTT	CCT	CCA	CCA	CCC	CCA	ATG	CCA	GGG	TCA	GGA	TTA	GGA	CCA	GGA	626
Phe	Leu	Pro	Pro	Pro	Pro	Pro	Met	Pro	Gly	Ser	Gly	Leu	Gly	Pro	Gly	
		190					195					200				
AAG	CCA	GGT	CTA	AAA	TTC	AAC	GGC	CCG	CCG	CCG	CCG	CCT	CCA	CTA	CCC	674
Lys	Pro	Gly	Leu	Lys	Phe	Asn	Gly	Pro	Pro	Pro	Pro	Pro	Pro	Pro	Leu	Pro
	205					210					215					



CCT	CCC	CCC	TTC	CTG	CCG	TGC	TGG	ATG	CCC	CCG	TTC	CCT	TCA	GGA	CCA	722
Pro	Pro	Pro	Phe	Leu	Pro	Cys	Trp	Met	Pro	Pro	Phe	Pro	Ser	Gly	Pro	
220					225					230					235	
CCA	ATA	ATC	CCG	CCA	CCC	CCT	CCC	ATC	TCT	CCC	GAC	TGT	CTG	GAT	GAC	770
Pro	Ile	Ile	Pro	Pro	Pro	Pro	Pro	Ile	Ser	Pro	Asp	Cys	Leu	Asp	Asp	
				240					245					250		
ACT	GAT	GCC	CTG	GGC	AGT	ATG	CTA	ATC	TCT	TGG	TAC	ATG	AGT	GGC	TAC	818
Thr	Asp	Ala	Leu	Gly	Ser	Met	Leu	Ile	Ser	Trp	Tyr	Met	Ser	Gly	Tyr	
			255					260					265			
CAC	ACT	GGC	TAC	TAT	ATG	GGT	TTC	AGA	CAA	AAT	AAA	AAA	GAA	GGA	AAG	866
His	Thr	Gly	Tyr	Tyr	Met	Gly	Phe	Arg	Gln	Asn	Lys	Lys	Glu	Gly	Lys	
		270					275					280				
TGC	TCA	CAT	ACA	AAT	TAAG											885
Cys	Ser	His	Thr	Asn												
285																

## (2) INFORMATION FOR SEQ ID NO:21:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 288 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: protein

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

Met	Ala	Met	Gly	Ser	Gly	Gly	Ala	Gly	Ser	Glu	Gln	Glu	Asp	Thr	Val	
1				5					10					15		
Leu	Phe	Arg	Arg	Gly	Thr	Gly	Gln	Ser	Asp	Asp	Ser	Asp	Ile	Trp	Asp	
			20					25					30			
Asp	Thr	Ala	Leu	Ile	Lys	Ala	Tyr	Asp	Lys	Ala	Val	Ala	Ser	Phe	Lys	
		35				40					45					
His	Ala	Leu	Lys	Asn	Gly	Asp	Ile	Cys	Glu	Thr	Pro	Asp	Lys	Pro	Lys	
	50				55						60					
Gly	Thr	Ala	Arg	Arg	Lys	Pro	Ala	Lys	Lys	Asn	Lys	Ser	Gln	Lys	Lys	
65					70					75				80		
Asn	Ala	Thr	Thr	Pro	Leu	Lys	Gln	Trp	Lys	Val	Gly	Asp	Lys	Cys	Ser	
				85					90					95		
Ala	Val	Trp	Ser	Glu	Asp	Gly	Cys	Ile	Tyr	Pro	Ala	Thr	Ile	Thr	Ser	
			100				105						110			

Ile	Asp	Phe	Lys	Arg	Glu	Thr	Cys	Val	Val	Val	Tyr	Thr	Gly	Tyr	Gly
		115					120					125			
Asn	Arg	Glu	Glu	Gln	Asn	Leu	Ser	Asp	Leu	Leu	Ser	Pro	Thr	Cys	Glu
	130					135					140				
Val	Ala	Asn	Ser	Thr	Glu	Gln	Asn	Thr	Gln	Glu	Asn	Glu	Ser	Gln	Val
145					150					155					160
Ser	Thr	Asp	Asp	Ser	Glu	His	Ser	Ser	Arg	Ser	Leu	Arg	Ser	Lys	Ala
				165					170					175	
His	Ser	Lys	Ser	Lys	Ala	Ala	Pro	Trp	Thr	Ser	Phe	Leu	Pro	Pro	Pro
			180					185					190		
Pro	Pro	Met	Pro	Gly	Ser	Gly	Leu	Gly	Pro	Gly	Lys	Pro	Gly	Leu	Lys
		195					200					205			
Phe	Asn	Gly	Pro	Pro	Pro	Pro	Pro	Pro	Leu	Pro	Pro	Pro	Pro	Phe	Leu
	210					215					220				
Pro	Cys	Trp	Met	Pro	Pro	Phe	Pro	Ser	Gly	Pro	Pro	Ile	Ile	Pro	Pro
225				230						235					240
Pro	Pro	Pro	Ile	Ser	Pro	Asp	Cys	Leu	Asp	Asp	Thr	Asp	Ala	Leu	Gly
				245					250					255	
Ser	Met	Leu	Ile	Ser	Trp	Tyr	Met	Ser	Gly	Tyr	His	Thr	Gly	Tyr	Tyr
			260					265					270		
Met	Gly	Phe	Arg	Gln	Asn	Lys	Lys	Glu	Gly	Lys	Cys	Ser	His	Thr	Asn
		275					280					285			

## (2) INFORMATION FOR SEQ ID NO:22:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 3246 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

CCTCCCGGGC ACCGTACTGT TCCGCTCCCA GAAGCCCCGG GCGCCGGAAG TCGTCACTCT  
 TAAGAAGGGA CGGGGCCCCA CGCTGCGCAC CCGCGGGTTT GCTATGGCGA TGAGCAGCGG

60

120

CGGCAGTGGT	GGCGGCGTCC	CGGAGCAGGA	GGATTCCGTG	CTGTTCCGGC	GCGGCACAGG	180
CAGGTGAGGT	CGCAGCCAGT	GCAGTCTCCC	TATTAGCGCT	CTCAGCACCC	TTCTTCCGGC	240
CCAACTCTCC	TTCCGCAGTA	ATTTTGTAT	GTGTGGATTA	AGATGACTCT	TGGTACTAAC	300
ATACATTTTC	TGATTAAACC	TATCTGACAT	GAGTTGTTTT	TATTTCTTAC	CCTTTCCAGA	360
GCGATGATTC	TGACATTTGG	GATGATACAG	CACTGATAAA	AGCATATGAT	AAAGCTGTGG	420
CTTCATTTAA	GGTATGAAAT	GCTTGTTAGT	CGTTTTCTTA	TTTTCTCGTT	ATTCATTTGG	480
AAAGGAATTG	ATAACATACG	ATAAAGTGTT	AAGTGCTTTC	TGAGGTGACG	GAGCCTTGAG	540
ACTAGCTTAT	AGTAGTAACT	GGGTTATGTC	GTGACTTTTA	TTCTGTGCAC	CACCCTGTAA	600
CATGTACATT	TTTATTCCTA	TTTTTCGTAGC	ATGCTCTAAA	GAATGGTGAC	ATTTGTGAAA	660
CTTGGGGTAA	ACCAAAAACC	ACACCTAAAA	GAAAACCTGC	TAAGAAGAAT	AAAAGCCAAA	720
AGAAGAATAC	TGCAGCTTCC	TTACAACAGG	TTATTTTAAA	ATGTTGAGGA	TTTAACTTCA	780
AAGGATGTCT	CATTAGTCCT	TATTTAATAG	TGTAAAATGT	CTTTAACTGC	AGGTCGATCA	840
AAACGAGATG	ATAGTTTGCC	CTCTTCAAAA	GAAATGTGTG	CATGTATATA	TCTTTGATTT	900
CTTTTGTAGT	GGAAAGTTGG	GGACAAATGT	TCTGCCATTT	GGTCAGAAGA	CGGTTGCATT	960
TACGCAGCTA	CCATTGCTTC	AATTGATTTT	AAGAGAGAAA	CCTGTGTTGT	GGTTTACACT	1020
GGATATGGAA	ATAGAGAGGA	GCAAAATCTG	TCCGATCTAC	TTTCCCCAAT	CTGTGAAGTA	1080
GCTATAATA	TAGAACAGAA	TGCTCAAGAG	GTAAGGATAC	AAAAAAAAAA	AAATTCAATT	1140
TCTGGAAGCA	GAGACTAGAT	GAGAAACTGT	TAAACAGTAT	ACAACCGAGG	CATTAATTTT	1200
TTCTTAATCA	CACCCTTATA	ACAAAAACCT	GCATATTTTT	TCTTTTAAA	GAATGAAAAT	1260
GAAAGCCAAG	TTTCAACAGA	TGAAAGTGAG	AACTCCAGGT	CTCCTGGAAA	TAAATCAGAT	1320
AACATCAAGC	CCAAATCTGC	TCCATGGAAC	TCTTTTCTCC	CTCCACCACC	CCCCATGCCA	1380
GGGCCAAGAC	TGGGACCAGG	AAAGGTAAAC	CTTCTATGAA	AGTTTTCAG	AAAATAGTTA	1440
ATGTCGGGAC	ATTTAACCTC	TCTGTAACT	AATTTGTAGC	TCTCCAATAT	TCTGGGTAAT	1500
TATTTTTATC	CTTTTGGTTT	TGAGTCCTTT	TTATTCCTAT	CATATTGAAA	TTGGTAAGTT	1560
AATTTTCCTT	TGAAATATTC	CTTATAGCCA	GGTCTAAAAT	TCAATGGCCC	ACCACCGCCA	1620
CCGCCACCAC	CACCACCCCA	CTTACTATCA	TGCTGGCTGC	CTCCATTTCC	TTCTGGACCA	1680

CCAGTAAGTA	AAAAAGAGTA	TAGGTTAGAT	TTTGCTTTCA	CATACAATTT	GATAATAGAC	1740
TTTACTTTTT	GTTTACTGGA	TATAACAAT	ATCTTTTCT	GTCTCCAGAT	AATTCCCCCA	1800
CCACCTCCCA	TATGTCCAGA	TTCTCTTGAT	GATGCTGATG	CTTTGGGAAG	TATGTTAATT	1860
TCATGGTACA	TGAGTGGCTA	TCATACTGGC	TATTATATGG	TAAGTAATCA	CTCAGCATCT	1920
TTTCCTGACA	ATTTTTTTGT	AGTTATGTGA	CTTTGTTTGG	TAAATTTATA	AAATACTACT	1980
CTGCAGCCTA	ATAATTGTTT	TCTTTGGGAT	AACTTTTAAA	GTACATTAAA	AGACTATCAA	2040
CTTAATTTCT	GATCATATTT	TGTTGAATAA	AATAAGTAAA	ATGTCTTG TG	AAACAAAATG	2100
CTTTTAAACA	TCCATATAAA	GCTATCTATA	TATAGCTATC	TATGTCTATA	TAGCTATTTT	2160
TTTTAACTTC	CTTTTATTTT	CCTTACAGGG	TTTCAGACAA	AATCAAAAAG	AAGGAAGGTG	2220
CTCACATTCC	TTAAATTAAG	GAGTAAGTCT	GCCAGCATTA	TGAAAGTGAA	TCTTACTTTT	2280
GTAAAACTTT	ATGGTTTGTG	GAAAACAAAT	GTTTTTGAAC	AGTTAAAAAG	TTCAGATGTT	2340
AAAAAGTTGA	AAGGTTAATG	TAAAACAATC	AATATTAAAG	AATTTTGATG	CCAAACTAT	2400
TAGATAAAAG	GTTAATCTAC	ATCCCTACTA	GAATTCTCAT	ACTTAACTGG	TTGGTTATGT	2460
GGAAGAAACA	TACTTTCACA	ATAAAGAGCT	TTAGGATATG	ATGCCATTTT	ATATCACTAG	2520
TAGGCAGACC	AGCAGACTTT	TTTTTATTGT	GATATGGGAT	AACCTAGGCA	TACTGCACTG	2580
TAGACTCTGA	CATATGAAGT	GCTCTAGTCA	AGTTTAACTG	GTGTCCACAG	AGGACATGGT	2640
TTAACTGGAA	TTCGTCAAGC	CTCTGGTTCT	AATTTCTCAT	TTGCAGGAAA	TGCTGGCATA	2700
GAGCAGCACT	AAATGACACC	ACTAAAGAAA	CGATCAGACA	GATCTGGAAT	GTGAAGCGTT	2760
ATAGAAGATA	ACTGGCCTCA	TTTCTTCAAA	ATATCAAGTG	TTGGGAAAGA	AAAAAGGAAG	2820
TGGAATGGGT	AACTCTTCTT	GATTAAAAGT	TATGTAATAA	CCAAATGCAA	TGTGAAATAT	2880
TTTACTGGAC	TCTTTTGAAA	AACCATCTAG	TAAAAGACTG	GGGTGGGGGT	GGGAGGCCAG	2940
CACGGTGGTG	AGGCAGTTGA	GAAAATTTGA	ATGTGGATTA	GATTTTGAAT	GATATTGGAT	3000
AATTATTGGT	AATTTTATGG	CCTGTGAGAA	GGGTGTTGTA	GTTTATAAAA	GACTGTCTTA	3060
ATTTGCATAC	TTAAGCATTT	AGGAATGAAG	TGTTAGAGTG	TCTTAAAATG	TTTCAAATGG	3120
TTTAACAAAA	TGTATGTGAG	GCGTATGTGG	CAAAATGTTA	CAGAATCTAA	CTGGTGGACA	3180
TGGCTGTTCA	TTGTACTGTT	TTTTTCTATC	TTCTATATGT	TTAAAAGTAT	ATAATAAAAA	3240

TATTTA

3246

## (2) INFORMATION FOR SEQ ID NO:23:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 637 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: DNA (genomic)

## (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

GATCTGCCTT CCTTCCTGCC CCCATGTTTG TCTTTCCTTG TTTGTCTTTA TATAGATCAA	60
GCAGGTTTTA AATTCCTAGT AGGAGCTTAC ATTTACTTTT CCAAGGGGGA GGGGGAATAA	120
ATATCTACAC ACACACACAC ACACACACCA CACTGGAGTT CGAGACGAGG CCTAAGCAAC	180
ATGCCGAAAC CCCGTCTCTA CTAAATACAA AAAATAGCTG AGCTTGGTGG CGCACGCCTA	240
TAGTCCTAGC TACTGGGGAG GCTGAGGTGG GAGGATCGCT TGAGCCCAAG AAGTCGAGGC	300
TGCAGTGAGC CGAGATCGCG CCGCTGCACT CCAGCCTGAG CGACAGGGCG AGGCTCTGTC	360
TCAAACAAA CAAACAAAAA AAAAAAGGAA AGGAAATATA ACACAGTGAA ATGAAAGGAT	420
TGAGAGAAAT GAAAAATATA CACGCCACAA ATGTGGGAGG GCGATAACCA CTCGTAGAAA	480
GCGTGAGAAG TTAACAAG CGGTCCTCCC GGGCACCGTA CTGTTCCGCT CCCAGAAGCC	540
CCGGGCGCCG GAAGTCGTCA CTCTTAAGAA GGGACGGGGC CCCACGCTGC GCACCCGCGG	600
GTTTGCTATG GCGATGAGCA GCGGCGGCAG TGGTGGC	637

## (2) INFORMATION FOR SEQ ID NO:24:

## (i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

## (ii) MOLECULE TYPE: other nucleic acid

## (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

AGGGCGAGGC TCTGTCTCA

19

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

CGGGAGGACC GCTTG TAGT

19

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 19 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

GCCGGAAGTC GTC ACTCTT

19

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

GGGTGCTGAG AGCGCTAATA

20

(2) INFORMATION FOR SEQ ID NO:28:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

TGTTGGATT AAGATGACTC

20

(2) INFORMATION FOR SEQ ID NO:29:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 19 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

CACTTTATCG TATGTTATC

19

(2) INFORMATION FOR SEQ ID NO:30:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTGTGCACCA CCCTGTAACA TG

22

(2) INFORMATION FOR SEQ ID NO:31:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 19 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

AAGGACTAAT GAGACATCC

19

(2) INFORMATION FOR SEQ ID NO:32:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear
- (ii) MOLECULE TYPE: other nucleic acid
- (iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:



CGAGATGATA GTTTGCCCTC

20

(2) INFORMATION FOR SEQ ID NO:33:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

AGCTACTTCA CAGATTGGGG AAAG

24

(2) INFORMATION FOR SEQ ID NO:34:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

CTCATCTAGT CTCTGCTTCC

20

(2) INFORMATION FOR SEQ ID NO:35:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TGGATATGGA AATAGAGAGG GAGC

24

(2) INFORMATION FOR SEQ ID NO:36:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 22 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CACCCTTATA ACAAAAACCT GC

22

(2) INFORMATION FOR SEQ ID NO:37:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 23 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

GAGAAAGGAG TTCCATGGAG CAG

23

(2) INFORMATION FOR SEQ ID NO:38:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

GAGAGGTTAA ATGTCCCGAC

20

(2) INFORMATION FOR SEQ ID NO:39:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 22 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

GTGAGAACTC CAGGTCTCCT GG

22

(2) INFORMATION FOR SEQ ID NO:40:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 20 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

TGAGTCTGTT TGAATTCAGG

20

(2) INFORMATION FOR SEQ ID NO:41:

(i) SEQUENCE CHARACTERISTICS:

(iii) HYPOTHETICAL: NO

GAAGGAAATG GAGGCAGCCA GC

22

```

(i) SEQUENCE CHARACTERISTICS:
    (A) LENGTH: 21 base pairs
    (B) TYPE: nucleic acid
    (C) STRANDEDNESS: single
    (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

```

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TTTCTACCCA TTAGAATCTG G

21

(i) SEQUENCE CHARACTERISTICS:  
 (A) LENGTH: 24 base pairs  
 (B) TYPE: nucleic acid  
 (C) STRANDEDNESS: single  
 (D) TOPOLOGY: linear

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

CCCCACTTAC TATCATGCTG GCTG

24

(2) INFORMATION FOR SEQ ID NO:44:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

CCAGACTTTA CTTTTTGTTT ACTG

24

(2) INFORMATION FOR SEQ ID NO:45:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

ATAGCCACTC ATGTACCATG A

21

(2) INFORMATION FOR SEQ ID NO:46:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

AAGAGTAATT TAAGCCTCAG ACAG

24

(2) INFORMATION FOR SEQ ID NO:47:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

CTGCCATATG TCCAGATTCT CTTG

24

(2) INFORMATION FOR SEQ ID NO:48:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 25 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

AGACTATCAA CTTAATTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:49:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TAAGGAATGT GAGCACCTTC CTTC

24

(2) INFORMATION FOR SEQ ID NO:50:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 25 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGACTATCAA CTTAATTCT GATCA

25

(2) INFORMATION FOR SEQ ID NO:51:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 24 base pairs

(B) TYPE: nucleic acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GTAAGATTCA CTTTCATAAT GCTG

24

(2) INFORMATION FOR SEQ ID NO:52:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 21 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

CTTTATGGTT TGTGGAAAAC A

21

(2) INFORMATION FOR SEQ ID NO:53:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 20 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

GGCATCATAT CCTAAAGCTC

20

(2) INFORMATION FOR SEQ ID NO:54:

- (i) SEQUENCE CHARACTERISTICS:
  - (A) LENGTH: 23 base pairs
  - (B) TYPE: nucleic acid
  - (C) STRANDEDNESS: single
  - (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:



GTAATAACCA AATGCAATGT GAA

23

(2) INFORMATION FOR SEQ ID NO:55:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

CTACAACACC CTTCTCACAG

20

(2) INFORMATION FOR SEQ ID NO:56:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 20 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:56:

GGTGTCCACA GAGGACATGG

20

(2) INFORMATION FOR SEQ ID NO:57:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 24 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: other nucleic acid

(iii) HYPOTHETICAL: NO

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:57:

AAGAGTTAAC CCATTCCAGC TTCC

24

CCATTCCAGC TTCC